

060231-0309

Sub
G1

- (i) APPLICANT: Inouye, Sumiko
Hsu, Mei-Yin
Eagle, Susan
Inouye, Masayori

(ii) TITLE OF INVENTION: Prokaryotic Reverse Transcriptase

(iii) NUMBER OF SEQUENCES: 45

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Weiser & Associates
(B) STREET: 230 South Fifteenth Street, Suite 500
(C) CITY: Philadelphia
(D) STATE: Pennsylvania
(E) COUNTRY: U.S.A.
(F) ZIP: 19102

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: US 08/269,118
(B) FILING DATE: 30-JUN-1994
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Weiser, Gerard J.
(B) REGISTRATION NUMBER: 19,763
(C) REFERENCE/DOCKET NUMBER: 377.5888P

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 215-875-8383
(B) TELEFAX: 215-875-8394

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2176 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 640..2094

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCATCCGCGC GGACACCCCC TCCTACGTGC CCCCCGACGC GGAGAGCGGC GTGGAGACGG 60
TGTACCGCGT TTCCCTGGAT GGTACACCTGG TGGCGGTGGA GTGGGGCCCCG CGCACGGGCT 120
CGCCGCGTCA CCAGCGGCTC TGGTTCGACT CGGATGCGGA AGCCCCCGGA GCCTACTTCG 180
CGCGCCTCGA GAAGTTGGCG GCTGACGGCT ACATCGACGC GGCTTCGGCA TTGGTCTAAA 240
CCCTTCAACC ACGGCTCGGC CGCCACGCGC GGCCGGCAGG ACAGGTGCGA CGAACAGACG 300
ACGACGTGCG CTTACGCGC GAGCAGCCGA GAGAGGTCCG GAGTGCATCA GCCTGAGCGC 360
CTCGAGCGGC GGAGCGGCGT TCGCCGCTC CGGTTGGAAT GCAGGACACT CTCCGCAAGG 420
TAGCCTGTTT TTGGCTCTCT CCCTCCTAGG CACTACGGCC AGGGTGGGTA GCGGAGCCAA 480
CGACGCCACC GCCGTTTACC CACCCCGGCC GTAGTGCCTA GGAGGGGAGA GCCGGTGAGG 540
CTACCGTGCC CCAGGTAAGA TGGTGGTGCT TTCCCGGCCT CCGTCGACTG CTCGCGCCAT 600
GTCCCGTCTT CCATCGCCGC GCCCGCCCAA GGTGCAGAC ATG ACC GCC AGG CTG 654
Met Thr Ala Arg Leu
1 5
GAC CCG TTC GTC CCC GCA GCT TCG CCG CAG GCC GTG CCC ACG CCC GAG 702
Asp Pro Phe Val Pro Ala Ala Ser Pro Gln Ala Val Pro Thr Pro Glu
10 15 20
CTC ACC GCT CCG TCG TCA GAC GCG GCC GCG AAG CGT GAA GCC CGC CGG 750
Leu Thr Ala Pro Ser Ser Asp Ala Ala Lys Arg Glu Ala Arg Arg
25 30 35
CTC GCG CAC GAA GCG TTG CTC GTC CGC GCG AAG GCC ATC GAC GAA GCG 798
Leu Ala His Glu Ala Leu Leu Val Arg Ala Lys Ala Ile Asp Glu Ala
40 45 50
GGC GGC GCC GAC GAC TGG GTG CAG GCG CAG CTC GTC TCC AAG GGG CTC 846
Gly Gly Ala Asp Asp Trp Val Gln Ala Gln Leu Val Ser Lys Gly Leu
55 60 65
GCG GTC GAG GAC CTG GAC TTC TCC AGC GCC TCC GAG AAG GAC AAG AAG 894
Ala Val Glu Asp Leu Asp Phe Ser Ser Ala Ser Glu Lys Asp Lys Lys
70 75 80 85
GCC TGG AAG GAG AAG AAG AAG GCC GAG GCC ACC GAG CGC CGC GCG CTG 942
Ala Trp Lys Glu Lys Lys Lys Ala Glu Ala Thr Glu Arg Arg Ala Leu
90 95 100
AAG CGT CAG GCG CAC GAG GCG TGG AAG GCC ACG CAC GTG GGC CAC CTG 990

Lys	Arg	Gln	Ala	His	Glu	Ala	Trp	Lys	Ala	Thr	His	Val	Gly	His	Leu	
			105					110					115			
GGC	GCG	GCC	GTG	CAC	TGG	GCG	GAG	GAC	CGC	CTG	GCC	GAC	GCG	TTC	GAC	1038
Gly	Ala	Gly	Val	His	Trp	Ala	Glu	Asp	Arg	Leu	Ala	Asp	Ala	Phe	Asp	
		120					125					130				
GTG	CCC	CAC	CGC	GAG	GAG	CGC	GCC	CGG	GCC	AAC	GGC	CTG	ACG	GAG	CTG	1086
Val	Pro	His	Arg	Glu	Glu	Arg	Ala	Arg	Ala	Asn	Gly	Leu	Thr	Glu	Leu	
	135					140					145					
GAC	TCC	GCG	GAG	GCG	CTG	GCC	AAG	GCG	CTG	GGG	CTG	AGC	GTC	TCC	AAG	1134
Asp	Ser	Ala	Glu	Ala	Leu	Ala	Lys	Ala	Leu	Gly	Leu	Ser	Val	Ser	Lys	
150					155					160					165	
CTC	CGC	TGG	TTC	GCG	TTC	CAC	CGG	GAG	GTC	GAC	ACG	GCC	ACG	CAC	TAC	1182
Leu	Arg	Trp	Phe	Ala	Phe	His	Arg	Glu	Val	Asp	Thr	Ala	Thr	His	Tyr	
				170					175					180		
GTG	AGC	TGG	ACC	ATT	CCG	AAG	CGG	GAC	GGC	AGC	AAG	CGC	ACG	ATT	ACG	1230
Val	Ser	Trp	Thr	Ile	Pro	Lys	Arg	Asp	Gly	Ser	Lys	Arg	Thr	Ile	Thr	
			185					190					195			
TCC	CCC	AAG	CCT	GAG	CTG	AAG	GCA	GCG	CAG	CGC	TGG	GTG	CTG	TCC	AAC	1278
Ser	Pro	Lys	Pro	Glu	Leu	Lys	Ala	Ala	Gln	Arg	Trp	Val	Leu	Ser	Asn	
		200					205					210				
GTC	GTG	GAG	CGG	CTG	CCG	GTC	CAC	GGC	GCC	GCC	CAC	GGC	TTC	GTG	GCG	1326
Val	Val	Glu	Arg	Leu	Pro	Val	His	Gly	Ala	Ala	His	Gly	Phe	Val	Ala	
	215					220					225					
GGA	CGC	TCC	ATC	CTC	ACC	AAC	GCG	CTG	GCC	CAC	CAG	GGC	GCG	GAC	GTC	1374
Gly	Arg	Ser	Ile	Leu	Thr	Asn	Ala	Leu	Ala	His	Gln	Gly	Ala	Asp	Val	
230					235					240					245	
GTG	GTC	AAG	GTG	GAC	CTC	AAG	GAC	TTC	TTC	CCC	TCC	GTC	ACC	TGG	CGC	1422
Val	Val	Lys	Val	Asp	Leu	Lys	Asp	Phe	Phe	Pro	Ser	Val	Thr	Trp	Arg	
				250				255						260		
CGG	GTG	AAG	GGC	CTG	TTG	CGC	AAG	GGC	GGC	CTG	CGG	GAG	GGC	ACG	TCC	1470
Arg	Val	Lys	Gly	Leu	Leu	Arg	Lys	Gly	Gly	Leu	Arg	Glu	Gly	Thr	Ser	
			265					270					275			
ACG	CTG	CTG	TCC	CTC	CTC	TCC	ACG	GAA	GCG	CCG	CGG	GAG	GCG	GTC	CAG	1518
Thr	Leu	Leu	Ser	Leu	Leu	Ser	Thr	Glu	Ala	Pro	Arg	Glu	Ala	Val	Gln	
		280					285					290				
TTC	CGC	GGC	AAG	CTC	CTG	CAC	GTC	GCC	AAG	GGC	CCG	CGC	GCC	CTG	CCC	1566
Phe	Arg	Gly	Lys	Leu	Leu	His	Val	Ala	Lys	Gly	Pro	Arg	Ala	Leu	Pro	
	295					300					305					
CAG	GGC	GCC	CCC	ACG	TCG	CCC	GGC	ATC	ACC	AAC	GCG	CTC	TGC	CTG	AAG	1614
Gln	Gly	Ala	Pro	Thr	Ser	Pro	Gly	Ile	Thr	Asn	Ala	Leu	Cys	Leu	Lys	
310					315					320					325	

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

CTC	GAC	AAG	CGG	CTG	TCC	GCC	CTC	GCG	AAG	CGG	CTG	GGC	TTC	ACC	TAC	1662
Leu	Asp	Lys	Arg	Leu	Ser	Ala	Leu	Ala	Lys	Arg	Leu	Gly	Phe	Thr	Tyr	
			330						335					340		
ACG	CGC	TAC	GCG	GAC	GAC	CTG	ACC	TTC	TCC	TGG	ACG	AAG	GCG	AAG	CAG	1710
Thr	Arg	Tyr	Ala	Asp	Asp	Leu	Thr	Phe	Ser	Trp	Thr	Lys	Ala	Lys	Gln	
			345					350					355			
CCC	AAG	CCG	CGG	CGG	ACG	CAG	CGT	CCC	CCC	GTC	GCG	GTC	CTC	CTG	TCT	1758
Pro	Lys	Pro	Arg	Arg	Thr	Gln	Arg	Pro	Pro	Val	Ala	Val	Leu	Leu	Ser	
		360					365					370				
CGC	GTC	CAG	GAA	GTG	GTG	GAG	GCG	GAG	GGC	TTC	CGC	GTG	CAC	CCG	GAC	1806
Arg	Val	Gln	Glu	Val	Val	Glu	Ala	Glu	Gly	Phe	Arg	Val	His	Pro	Asp	
	375					380					385					
AAG	ACG	CGC	GTC	GCC	CGC	AAG	GGC	ACG	CGG	CAG	CGG	GTC	ACC	GGG	CTC	1854
Lys	Thr	Arg	Val	Ala	Arg	Lys	Gly	Thr	Arg	Gln	Arg	Val	Thr	Gly	Leu	
					395					400					405	
GTC	GTG	AAT	GCG	GCG	GGC	AAG	GAC	GCG	CCC	GCG	GCC	CGA	GTC	CCG	CGC	1902
Val	Val	Asn	Ala	Ala	Gly	Lys	Asp	Ala	Pro	Ala	Ala	Arg	Val	Pro	Arg	
				410					415					420		
GAC	GTC	GTC	CGC	CAG	CTC	CGC	GCC	GCC	ATC	CAC	AAC	CGG	AAG	AAG	GGC	1950
Asp	Val	Val	Arg	Gln	Leu	Arg	Ala	Ala	Ile	His	Asn	Arg	Lys	Lys	Gly	
			425					430					435			
AAG	CCG	GGC	CGC	GAG	GGC	GAG	TCG	CTC	GAG	CAG	CTC	AAG	GGC	ATG	GCC	1998
Lys	Pro	Gly	Arg	Glu	Gly	Glu	Ser	Leu	Glu	Gln	Leu	Lys	Gly	Met	Ala	
		440					445					450				
GCC	TTC	ATC	CAC	ATG	ACG	GAC	CCG	GCC	AAG	GGC	CGC	GCC	TTC	CTG	GCT	2046
Ala	Phe	Ile	His	Met	Thr	Asp	Pro	Ala	Lys	Gly	Arg	Ala	Phe	Leu	Ala	
	455					460					465					
CAG	CTC	ACG	GAG	CTC	GAG	TCC	ACG	GCG	AGC	GCC	GCT	CCG	CAG	GCG	GAG	2094
Gln	Leu	Thr	Glu	Leu	Glu	Ser	Thr	Ala	Ser	Ala	Ala	Pro	Gln	Ala	Glu	
	470				475				480						485	
TGACGCTCAG CGCGCGTCCG TCGCCGACGT GCCGCGCGCC AGCAACGCCG CATTCAGCAA																2154
CTCCGTCAGC CGGCGCGGGT AC																2176

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro
1 5 10 15
Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met
20 25 30
Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn
35 40 45
Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys
50 55 60
Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu
65 70 75 80
Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser
85 90 95
Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp
100 105 110
Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn
115 120 125
Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp
130 135 140
Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu
145 150 155 160
Pro Phe Lys Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp
165 170 175
Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys
180 185 190
Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro
195 200 205
Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu
210 215 220
Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys
225 230 235 240
Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn
245 250 255
Trp Ala Ser Gln Ile Tyr Pro
260

09303031-00000000

RECEIVED - 03/06/2011

(A) LENGTH: 263 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg 1	Pro	Trp	Ala 5	Arg	Thr	Pro	Pro	Lys 10	Ala	Pro	Arg	Asn	Gln 15	Pro	Val
Pro	Phe	Lys 20	Pro	Glu	Arg	Leu	Gln	Ala 25	Leu	Gln	His	Leu	Val 30	Arg	Lys
Ala	Leu 35	Glu	Ala	Gly	His	Ile	Glu 40	Pro	Tyr	Thr	Gly	Pro 45	Gly	Asn	Asn
Pro	Val 50	Phe	Pro	Val	Lys	Lys 55	Ala	Asn	Gly	Thr	Trp 60	Arg	Phe	Ile	His
Asp 65	Leu	Arg	Ala	Thr	Asn 70	Ser	Leu	Thr	Ile	Asp 75	Leu	Ser	Ser	Ser	Ser 80
Pro	Gly	Pro	Pro	Asp 85	Leu	Ser	Ser	Leu	Pro 90	Thr	Thr	Leu	Ala	His 95	Leu
Gln	Thr	Ile	Asp 100	Leu	Arg	Asp	Ala	Phe 105	Phe	Gln	Ile	Pro	Leu 110	Pro	Lys
Gln	Phe 115	Gln	Pro	Tyr	Phe	Ala	Phe 120	Thr	Val	Pro	Gln	Gln 125	Cys	Asn	Tyr
Gly 130	Pro	Gly	Thr	Arg	Tyr	Ala 135	Trp	Lys	Val	Leu	Pro 140	Gln	Gly	Phe	Lys
Asn 145	Ser	Pro	Thr	Leu	Phe 150	Glu	Met	Gln	Leu	Ala 155	His	Ile	Leu	Gln 160	Pro
Ile	Arg	Gln	Ala	Phe 165	Pro	Gln	Cys	Thr	Ile 170	Leu	Gln	Tyr	Met	Asp 175	Asp
Ile	Leu	Leu 180	Ala	Ser	Pro	Ser	His	Glu 185	Asp	Leu	Leu	Leu 190	Leu	Ser	Glu
Ala	Thr 195	Met	Ala	Ser	Leu	Ile	Ser 200	His	Gly	Leu	Pro	Val 205	Ser	Glu	Asn
Lys 210	Thr	Gln	Gln	Thr	Pro	Gly 215	Thr	Ile	Lys	Phe	Leu 220	Gly	Gln	Ile	Ile

Ser Pro Asn His Leu Thr Tyr Asp Ala Val Pro Thr Val Pro Ile Arg
225 230 235 240

Ser Arg Trp Ala Leu Pro Glu Leu Gln Ala Leu Leu Gly Glu Ile Gln
245 250 255

Trp Val Ser Lys Gly Thr Pro
260

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 259 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asn Val Leu Tyr Arg Ile Gly Ser Asp Asn Gln Tyr Thr Gln Phe Thr
1 5 10 15

Ile Pro Lys Lys Gly Lys Gly Val Arg Thr Ile Ser Ala Pro Thr Asp
20 25 30

Arg Leu Lys Asp Ile Gln Arg Arg Ile Cys Asp Leu Leu Ser Asp Cys
35 40 45

Arg Asp Glu Ile Phe Ala Ile Arg Lys Ile Ser Asn Asn Tyr Ser Phe
50 55 60

Gly Phe Glu Arg Gly Lys Ser Ile Ile Leu Asn Ala Tyr Lys His Arg
65 70 75 80

Gly Lys Gln Ile Ile Leu Asn Ile Asp Leu Lys Asp Phe Phe Glu Ser
85 90 95

Phe Asn Phe Gly Arg Val Arg Gly Tyr Phe Leu Ser Asn Gln Asp Phe
100 105 110

Leu Leu Asn Pro Val Val Ala Thr Thr Leu Ala Lys Ala Ala Cys Tyr
115 120 125

Asn Gly Thr Leu Pro Gln Gly Ser Pro Cys Ser Pro Ile Ile Ser Asn
130 135 140

Leu Ile Cys Asn Ile Met Asp Met Arg Leu Ala Lys Leu Ala Lys Lys
145 150 155 160

Tyr Gly Cys Thr Tyr Ser Arg Tyr Ala Asp Asp Ile Thr Ile Ser Thr
165 170 175

000000001.030000000
Sue
G
cont

Asn Lys Asn Thr Phe Pro Leu Glu Met Ala Thr Val Gln Pro Glu Gly
 180 185 190
 Val Val Leu Gly Lys Val Leu Val Lys Glu Ile Glu Asn Ser Gly Phe
 195 200 205
 Glu Ile Asn Asp Ser Lys Thr Arg Leu Thr Tyr Lys Thr Ser Arg Gln
 210 215 220
 Glu Val Thr Gly Leu Thr Val Asn Arg Ile Val Asn Ile Asp Arg Cys
 225 230 235 240
 Tyr Tyr Lys Lys Thr Arg Ala Leu Ala His Ala Leu Tyr Arg Thr Gly
 245 250 255
 Glu Tyr Lys

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ala Phe His Arg Glu Val Asp Thr Ala Thr His Tyr Val Ser Trp Thr
 1 5 10 15
 Ile Pro Lys Arg Asp Gly Ser Lys Arg Thr Ile Thr Ser Pro Lys Pro
 20 25 30
 Glu Leu Lys Ala Ala Gln Arg Trp Val Leu Ser Asn Val Val Glu Arg
 35 40 45
 Leu Pro Val His Gly Ala Ala His Gly Phe Val Ala Gly Arg Ser Ile
 50 55 60
 Leu Thr Asn Ala Leu Ala His Gln Gly Ala Asp Val Val Val Lys Val
 65 70 75 80
 Asp Leu Lys Asp Phe Phe Pro Ser Val Thr Trp Arg Arg Val Lys Gly
 85 90 95
 Leu Leu Arg Lys Gly Gly Leu Arg Glu Gly Thr Ser Thr Leu Leu Ser
 100 105 110
 Leu Leu Ser Thr Glu Ala Pro Arg Glu Ala Val Gln Phe Arg Gly Lys
 115 120 125

CY
 G
 cont.

Arg Arg Ala Leu Lys Arg Gln Ala His Glu Ala Trp Lys Ala Thr His
85 90 95

Val Gly His Leu Gly Ala Gly Val His Trp Ala Glu Asp Arg Leu
100 105 110

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 110 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro Asp Pro Asp Met Thr Arg Val Thr Asn Ser Pro Ser Leu Gln Ala
1 5 10 15

His Leu Gln Ala Leu Tyr Leu Val Gln His Glu Val Trp Arg Pro Leu
20 25 30

Ala Ala Ala Tyr Gln Glu Gln Leu Asp Arg Pro Val Val Pro His Pro
35 40 45

Tyr Arg Val Gly Asp Thr Val Trp Val Arg Arg His Gln Thr Lys Asn
50 55 60

Leu Glu Pro Arg Trp Lys Gly Pro Tyr Thr Val Leu Leu Thr Thr Pro
65 70 75 80

Thr Ala Leu Lys Val Asp Gly Ile Ala Ala Trp Ile His Ala Ala His
85 90 95

Val Lys Ala Ala Asp Pro Gly Gly Gly Pro Ser Ser Arg Leu
100 105 110

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gly Lys Asp Ala Pro Ala Ala Arg Val Pro Arg Asp Val Val Arg Gln

RECEIVED - FEB 28 1981

Sub
Cont

1				5					10					15		
Leu	Arg	Ala	Ala	Ile	His	Asn	Arg	Lys	Lys	Gly	Lys	Pro	Gly	Arg	Glu	
			20					25					30			
Gly	Glu	Ser	Leu	Glu	Gln	Leu	Lys	Gly	Met	Ala	Ala	Phe	Ile	His	Met	
		35					40					45				
Thr	Asp	Pro	Ala	Lys	Gly	Arg	Ala	Phe	Leu	Ala	Gln	Leu	Thr	Glu	Leu	
	50					55					60					
Glu	Ser	Thr	Ala	Ser	Ala	Ala	Pro	Gln	Ala	Glu						
65					70					75						

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gly Lys Glu Gly His Ser Ala Arg Gln Cys Arg Ala Pro Arg Arg Gln
1 5 10 15

Gly Cys Trp Lys Cys Gly Lys / Pro Gly His Ile Met Thr Asn Cys Pro
20 25 30

Asp Arg Gln Ala Gly Phe Leu ~~Gly~~ Leu Gly Pro Trp Gly Lys Lys Pro
35 40 45

Arg Asn Phe Pro Val Ala Gln Val Pro Gln Gly Leu Thr Pro Thr Ala
50 55 60

Pro Pro
65

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gly Pro Arg Ala Leu Pro Gln Gly Ala Pro Thr Ser Pro Gly Ile Thr
 1 5 10 15
 Asn Ala Leu Cys Leu Lys Leu Asp Lys Arg Leu Ser Ala Leu Ala Lys
 20 25 30
 Arg Leu Gly Phe Thr Tyr Thr Arg Tyr Ala Asp Asp Leu Thr Phe Ser
 35 40 45
 Trp Thr Lys Ala Lys Gln Pro Lys Pro Arg Arg Thr Gln Arg Pro Pro
 50 55 60
 Val Ala Val Leu
 65

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 68 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Tyr Asn Gly Thr Leu Pro Gln Gly Ser Pro Cys Ser Pro Ile Ile Ser
 1 5 10 15
 Asn Leu Ile Cys Asn Ile Met Asp Met Arg Leu Ala Lys Leu Ala Lys
 20 25 30
 Lys Tyr Gly Cys Thr Tyr Ser Arg Tyr Ala Asp Asp Ile Thr Ile Ser
 35 40 45
 Thr Asn Lys Asn Thr Phe Pro Leu Glu Met Ala Thr Val Gln Pro Glu
 50 55 60
 Gly Val Val Leu
 65

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 68 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

RECORDED - 030382

Sub
 G-1
 Cont

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ile Ala Thr Asn Gly Val Pro Gln Gly Ala Ser Thr Ser Cys Gly Leu
1 5 10 15
Ala Thr Tyr Asn Val Leu Glu Leu Phe Leu Arg Tyr Asp Glu Leu Ile
20 25 30
Met Tyr Ala Asp Asp Gly Ile Leu Cys Arg Gln Asp Pro Ser Thr Pro
35 40 45
Asp Phe Ser Val Glu Glu Ala Gly Val Val Gln Glu Pro
50 55 60

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Tyr Glu Tyr Leu Arg Met Pro Phe Gly Leu Lys Asn Ala Pro Ala Thr
1 5 10 15
Phe Gln Arg Cys Met Asn Asp Ile Leu Arg Pro Leu Leu Asn Lys His
20 25 30
Cys Leu Val Tyr Leu Asp Asp Ile Ile Val Phe Ser Thr Ser Leu Asp
35 40 45
Glu His Leu Gln Ser Leu Gly Leu Val Phe Glu Lys Leu
50 55 60

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

030301-030307

[illegible]

060809 1. 030357Z

[illegible]

000000 1 - 030397

- 00000000 - 030397

[illegible]

06000001-030397

[illegible]

069807 : 030352

068000-1 - 030397

060809. 030357Z

060809. 030357Z

[illegible][illegible]

Trp Asn Ala Cys Val Tyr Val Lys Gln Val Ser Glu Gln Glu His Leu
20 25 30
Tyr Leu Leu Leu Tyr Val Asp Asp Met Leu Ile Ala Gly Lys Ser Lys
35 40 45
Ser Glu Ile Asn Lys Val Lys Glu Gln Leu Ser Met Glu Phe
50 55 60

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Ile Arg Leu Lys Lys Ser Leu Tyr Glu Leu Lys Gln Ser Gly Ala Asn
1 5 10 15
Trp Tyr Glu Glu Val Arg Gly Trp Ser Cys Val Phe Lys Asn Ser Gln
20 25 30
Val Thr Ile Cys Leu Phe Val Asp Asp Met Val Leu Phe Ser Lys Asn
35 40 45
Leu Asn Ser Asn Lys Arg Ile Ile Glu Lys Leu Lys Met Gln Tyr
50 55 60

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "The 2' position of this
nucleotide is linked to the 5' position of
nucleotide number 1 of SEQ ID NO: 25 of this
application."

(ix) FEATURE:

- (A) NAME/KEY: misc_binding
(B) LOCATION: 52..58

26E0E0-TE080800

030499-1

CACGCAUGUA GGCAGAUUUG UUGGUUGUGA AUCGCAACCA GUGGCCUUAU UGGCAGGA

58

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 67 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(A) NAME/KEY: misc_feature
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "The 5' position of this nucleotide is linked to the 2' position of nucleotide number 15 of SEQ ID NO: 24 of this application."

(A) NAME/KEY: misc_binding
(B) LOCATION: 61..67
(D) OTHER INFORMATION: /note= "This region can hydrogen bond to nucleotides 52-58 of SEQ ID NO: 24 of this application."

TCCTTCGCAC AGCACACCTG CCGTATAGCT CTGAATCAAG GATTTTAGGG AGGCGATTCC

60

67

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2423 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(A) NAME/KEY: CDS
(B) LOCATION: 418..2175

145					150					155					160			
GGA Gly	ACC Thr	CTC Leu	CCC Pro	CAA Gln 165	GGA Gly	AGT Ser	CCA Pro	TGT Cys	TCT Ser 170	CCT Pro	ATT Ile	ATC Ile	TCA Ser	AAT Asn 175	CTA Leu		945	
ATT Ile	TGC Cys	AAT Asn	ATT Ile 180	ATG Met	GAT Asp	ATG Met	AGA Arg	TTA Leu 185	GCT Ala	AAG Lys	CTG Leu	GCT Ala	AAA Lys 190	AAA Lys	TAT Tyr		993	
GGA Gly	TGT Cys	ACT Thr 195	TAT Tyr	AGC Ser	AGA Arg	TAT Tyr	GCT Ala 200	GAT Asp	GAT Asp	ATA Ile	ACA Thr	ATT Ile 205	TCT Ser	ACA Thr	AAT Asn		1041	
AAA Lys	AAT Asn 210	ACA Thr	TTT Phe	CCG Pro	TTA Leu	GAA Glu 215	ATG Met	GCT Ala	ACT Thr	GTG Val	CAA Gln 220	CCT Pro	GAA Glu	GGG Gly	GTT Val		1089	
GTT Val 225	TTG Leu	GGA Gly	AAA Lys	GTT Val	TTG Leu 230	GTA Val	AAA Lys	GAA Glu	ATA Ile	GAA Glu 235	AAC Asn	TCT Ser	GGA Gly	TTC Phe	GAA Glu 240		1137	
ATA Ile	AAT Asn	GAT Asp	TCA Ser 245	AAG Lys	ACT Thr	AGG Arg	CTT Leu	ACG Thr 250	TAT Tyr	AAG Lys	ACA Thr	TCA Ser	AGG Arg	CAA Gln 255	GAA Glu		1185	
GTA Val	ACG Thr	GGA Gly	CTT Leu 260	ACA Thr	GTT Val	AAC Asn	AGA Arg	ATC Ile 265	GTT Val	AAT Asn	ATT Ile	GAT Asp 270	AGA Arg	TGT Cys	TAT Tyr		1233	
TAT Tyr	AAA Lys	AAA Lys 275	ACT Thr	CGG Arg	GCG Ala	TTG Leu	GCA Ala 280	CAT His	GCT Ala	TTG Leu	TAT Tyr	CGT Arg 285	ACA Thr	GGT Gly	GAA Glu		1281	
TAT Tyr	AAA Lys 290	GTG Val	CCA Pro	GAT Asp	GAA Glu	AAT Asn 295	GGT Gly	GTT Val	TTA Leu	GTT Val	TCA Ser 300	GGA Gly	GGT Gly	CTG Leu	GAT Asp		1329	
AAA Lys 305	CTT Leu	GAG Glu	GGG Gly	ATG Met 310	TTT Phe	GGT Gly	TTT Phe	ATT Ile	GAT Asp	CAA Gln 315	GTT Val	GAT Asp	AAG Lys	TTT Phe	AAC Asn 320		1377	
AAT Asn	ATA Ile	AAG Lys	AAA Lys 325	AAA Lys	CTG Leu	AAC Asn	AAG Lys	CAA Gln	CCT Pro 330	GAT Asp	AGA Arg	TAT Tyr	GTA Val	TTG Leu 335	ACT Thr		1425	
AAT Asn	GCG Ala	ACT Thr 340	TTG Leu	CAT His	GGT Gly	TTT Phe	AAA Lys	TTA Leu 345	AAG Lys	TTG Leu	AAT Asn	GCG Ala 350	CGA Arg	GAA Glu	AAA Lys		1473	
GCA Ala	TAT Tyr	AGT Ser 355	AAA Lys	TTT Phe	ATT Ile	TAC Tyr 360	TAT Tyr	AAA Lys	TTT Phe	TTT Phe	CAT His 365	GGC Gly	AAC Asn	ACC Thr	TGT Cys		1521	

CCT Pro	ACG Thr 370	ATA Ile	ATT Ile	ACA Thr	GAA Glu	GGG Gly 375	AAG Lys	ACT Thr	GAT Asp	CGG Arg	ATA Ile 380	TAT Tyr	TTG Leu	AAG Lys	GCT Ala	1569
GCT Ala 385	TTG Leu	CAT His	TCT Ser	TTG Leu	GAG Glu 390	ACA Thr	TCA Ser	TAT Tyr	CCT Pro	GAG Glu 395	TTG Leu	TTT Phe	AGA Arg	GAA Glu	AAA Lys 400	1617
ACA Thr	GAT Asp	AGT Ser	AAA Lys	AAG Lys 405	AAA Lys	GAA Glu	ATA Ile	AAT Asn	CTT Leu 410	AAT Asn	ATA Ile	TTT Phe	AAA Lys	TCT Ser 415	AAT Asn	1665
GAA Glu	AAG Lys	ACC Thr	AAA Lys 420	TAT Tyr	TTT Phe	TTA Leu	GAT Asp	CTT Leu 425	TCT Ser	GGG Gly	GGA Gly	ACT Thr	GCA Ala 430	GAT Asp	CTG Leu	1713
AAA Lys	AAA Lys	TTT Phe 435	GTA Val	GAG Glu	CGT Arg	TAT Tyr	AAA Lys 440	AAT Asn	AAT Asn	TAT Tyr	GCT Ala	TCT Ser 445	TAT Tyr	TAT Tyr	GGT Gly	1761
TCT Ser	GTT Val 450	CCA Pro	AAA Lys	CAG Gln	CCA Pro	GTG Val 455	ATT Ile	ATG Met	GTT Val	CTT Leu	GAT Asp 460	AAT Asn	GAT Asp	ACA Thr	GGT Gly	1809
CCA Pro 465	AGC Ser	GAT Asp	TTA Leu	CTT Leu	AAT Asn 470	TTT Phe	CTG Leu	CGC Arg	AAT Asn	AAA Lys 475	GTT Val	AAA Lys	AGC Ser	TGC Cys	CCA Pro 480	1857
GAC Asp	GAT Asp	GTA Val	ACT Thr	GAA Glu 485	ATG Met	AGA Arg	AAG Lys	ATG Met	AAA Lys 490	TAT Tyr	ATT Ile	CAT His	GTT Val	TTC Phe 495	TAT Tyr	1905
AAT Asn	TTA Leu	TAT Tyr	ATA Ile 500	GTT Val	CTC Leu	ACA Thr	CCA Pro	TTG Leu 505	AGT Ser	CCT Pro	TCC Ser	GGC Gly	GAA Glu 510	CAA Gln	ACT Thr	1953
TCA Ser	ATG Met	GAG Glu 515	GAT Asp	CTT Leu	TTC Phe	CCT Pro	AAA Lys 520	GAT Asp	ATT Ile	TTA Leu	GAT Asp	ATC Ile 525	AAG Lys	ATT Ile	GAT Asp	2001
GGT Gly	AAG Lys 530	AAA Lys	TTC Phe	AAC Asn	AAA Lys	AAT Asn 535	AAT Asn	GAT Asp	GGA Gly	GAC Asp	TCA Ser 540	AAA Lys	ACG Thr	GAA Glu	TAT Tyr	2049
GGG Gly 545	AAG Lys	CAT His	ATT Ile	TTT Phe	TCC Ser 550	ATG Met	AGG Arg	GTT Val	GTT Val	AGA Arg 555	GAT Asp	AAA Lys	AAG Lys	CGG Arg	AAA Lys 560	2097
ATA Ile	GAT Asp	TTT Phe	AAG Lys	GCA Ala 565	TTT Phe	TGT Cys	TGT Cys	ATT Ile	TTT Phe 570	GAT Asp	GCT Ala	ATA Ile	AAA Lys	GAT Asp 575	ATA Ile	2145
AAG Lys	GAA Glu	CAT His	TAT Tyr	AAA Lys	TTA Leu	ATG Met	TTA Leu	AAT Asn	AGC Ser	TAATGAACAG CCCTAACGTT						2195

585

ATGAACGCTA	AGGCTGATT	TTCGTTAAAA	TTTATATGGT	TTGAATTGTA	ATATATTATC	2255
TTCAAGCCAT	TTATTTAATT	CCTGCATCCT	TTTCTGTAAG	GGTATTAATT	CGTTCCTCAC	2315
AAACACTAAA	CTCGCTTTTT	CCACATCCCC	AAACCCCCCT	AACATTATTC	GGCATAATCC	2375
CCATCATTTG	CGGTGGCACA	CGATGCGCTG	CCATCATGTC	ATCGCGGC		2423

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 546 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Val 1	Lys	Leu	Lys	Pro 5	Gly	Met	Asp	Gly	Pro 10	Lys	Val	Lys	Gln	Trp 15	Pro
Leu	Thr	Glu	Glu 20	Lys	Ile	Lys	Ala	Leu 25	Val	Glu	Ile	Cys	Thr 30	Glu	Met
Glu	Lys	Glu 35	Gly	Lys	Ile	Ser	Lys 40	Ile	Gly	Pro	Glu	Asn 45	Pro	Tyr	Asn
Thr	Pro 50	Val	Phe	Ala	Ile	Lys 55	Lys	Lys	Asp	Ser	Thr 60	Lys	Trp	Arg	Lys
Leu 65	Val	Asp	Phe	Arg	Glu 70	Leu	Asn	Lys	Arg	Thr 75	Gln	Asp	Phe	Trp	Glu 80
Val	Gln	Leu	Gly	Ile 85	Pro	His	Pro	Ala	Gly 90	Leu	Lys	Lys	Lys	Lys 95	Ser
Val	Thr	Val	Leu 100	Asp	Val	Gly	Asp	Ala 105	Tyr	Phe	Ser	Val	Pro 110	Leu	Asp
Glu	Asp	Phe 115	Arg	Lys	Tyr	Thr	Ala 120	Phe	Thr	Ile	Pro	Ser 125	Ile	Asn	Asn
Glu	Thr 130	Pro	Gly	Ile	Arg	Tyr 135	Gln	Tyr	Asn	Val	Leu 140	Pro	Gln	Gly	Trp
Lys 145	Gly	Ser	Pro	Ala	Ile 150	Phe	Gln	Ser	Ser	Met 155	Thr	Lys	Ile	Leu	Glu 160

SECRET

Gln	Phe	Gln	Pro	Tyr	Phe	Ala	Phe	Thr	Val	Pro	Gln	Gln	Cys	Asn	Tyr
		115					120					125			
Gly	Pro	Gly	Thr	Arg	Tyr	Ala	Trp	Lys	Val	Leu	Pro	Gln	Gly	Phe	Lys
	130					135					140				
Asn	Ser	Pro	Thr	Leu	Phe	Glu	Met	Gln	Leu	Ala	His	Ile	Leu	Gln	Pro
145					150					155					160
Ile	Arg	Gln	Ala	Phe	Pro	Gln	Cys	Thr	Ile	Leu	Gln	Tyr	Met	Asp	Asp
				165					170					175	
Ile	Leu	Leu	Ala	Ser	Pro	Ser	His	Glu	Asp	Leu	Leu	Leu	Leu	Ser	Glu
			180					185					190		
Ala	Thr	Met	Ala	Ser	Leu	Ile	Ser	His	Gly	Leu	Pro	Val	Ser	Glu	Asn
		195					200					205			
Lys	Thr	Gln	Gln	Thr	Pro	Gly	Thr	Ile	Lys	Phe	Leu	Gly	Gln	Ile	Ile
	210					215					220				
Ser	Pro	Asn	His	Leu	Thr	Tyr	Asp	Ala	Val	Pro	Thr	Val	Pro	Ile	Arg
225					230					235					240
Ser	Arg	Trp	Ala	Leu	Pro	Glu	Leu	Gln	Ala	Leu	Leu	Gly	Glu	Ile	Gln
				245					250					255	
Trp	Val	Ser	Lys	Gly	Thr	Pro	Thr	Leu	Arg	Gln	Pro	Leu	His	Ser	Leu
			260					265					270		
Tyr	Cys	Ala	Leu	Gln	Arg	His	Thr	Asp	Pro	Arg	Asp	Gln	Ile	Tyr	Leu
		275					280					285			
Asn	Pro	Ser	Gln	Val	Gln	Ser	Leu	Val	Gln	Leu	Arg	Gln	Ala	Leu	Ser
	290					295					300				
Gln	Asn	Cys	Arg	Ser	Arg	Leu	Val	Gln	Thr	Leu	Pro	Leu	Leu	Gly	Ala
305					310					315					320
Ile	Met	Leu	Thr	Leu	Thr	Gly	Thr	Thr	Thr	Val	Val	Phe	Gln	Ser	Lys
				325				330						335	
Glu	Gln	Trp	Pro	Leu	Val	Trp	Leu	His	Ala	Pro	Leu	Pro	His	Thr	Ser
			340					345					350		
Gln	Cys	Pro	Trp	Gly	Gln	Leu	Leu	Ala	Ser	Ala	Val	Leu	Leu	Leu	Asp
		355					360					365			
Lys	Tyr	Thr	Leu	Gln	Ser	Tyr	Gly	Leu	Leu	Cys	Gln	Thr	Ile	His	His
	370					375					380				
Asn	Ile	Ser	Thr	Gln	Thr	Phe	Asn	Gln	Phe	Ile	Gln	Thr	Ser	Asp	His
385					390					395					400

RECEIVED - DECEMBER 1967

Pro Ser Val Pro Ile Leu Leu His His Ser His Arg Phe Lys Asn Leu
405 410 415

Gly Ala Gln Thr Gly Glu Leu Trp Asn Thr Phe Leu Lys Thr Ala Ala
420 425 430

Pro Leu Ala Pro Val Lys Ala Leu Met Pro Val Phe Thr Leu Ser Pro
435 440 445

Val Ile Ile Asn Thr Ala Pro Cys Leu Phe Ser Asp Gly Ser Thr Ser
450 455 460

Arg Ala Ala Tyr Ile Leu Trp Asp Lys Gln Ile Leu Ser Gln Arg Ser
465 470 475 480

Phe Pro Leu Pro Pro Pro His Lys Ser Ala Gln Arg Ala Glu Leu Leu
485 490 495

Gly Leu Leu His Gly Leu Ser Ser Ala Arg Ser Trp Arg Cys Leu Asn
500 505 510

Ile Phe Leu Asp Ser Lys Tyr Leu Tyr His Tyr Leu Arg Thr Leu Ala
515 520 525

Leu Gly Thr Phe Gln Gly Arg Ser Ser Gln Ala Pro Phe Gln Ala Leu
530 535 540

Leu Pro Arg Leu Leu Ser Arg Lys Val Val Tyr Leu His His Val Arg
545 550 555 560

Ser His Thr Asn Leu Pro Asp Pro Ile Ser Arg Leu Asn Ala Leu Thr
565 570 575

Asp Ala

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 555 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Asn Val Leu Tyr Arg Ile Gly Ser Asp Asn Gln Tyr Thr Gln Phe Thr
1 5 10 15

Ile Pro Lys Lys Gly Lys Gly Val Arg Thr Ile Ser Ala Pro Thr Asp
20 25 30

0300331.030397

Arg	Leu	Lys	Asp	Ile	Gln	Arg	Arg	Ile	Cys	Asp	Leu	Leu	Ser	Asp	Cys
		35					40					45			
Arg	Asp	Glu	Ile	Phe	Ala	Ile	Arg	Lys	Ile	Ser	Asn	Asn	Tyr	Ser	Phe
	50					55					60				
Gly	Phe	Glu	Arg	Gly	Lys	Ser	Ile	Ile	Leu	Asn	Ala	Tyr	Lys	His	Arg
65					70					75					80
Gly	Lys	Gln	Ile	Ile	Leu	Asn	Ile	Asp	Leu	Lys	Asp	Phe	Phe	Glu	Ser
				85					90					95	
Phe	Asn	Phe	Gly	Arg	Val	Arg	Gly	Tyr	Phe	Leu	Ser	Asn	Gln	Asp	Phe
			100					105					110		
Leu	Leu	Asn	Pro	Val	Val	Ala	Thr	Thr	Leu	Ala	Lys	Ala	Ala	Cys	Tyr
		115					120					125			
Asn	Gly	Thr	Leu	Pro	Gln	Gly	Ser	Pro	Cys	Ser	Pro	Ile	Ile	Ser	Asn
	130					135					140				
Leu	Ile	Cys	Asn	Ile	Met	Asp	Met	Arg	Leu	Ala	Lys	Leu	Ala	Lys	Lys
145					150					155					160
Tyr	Gly	Cys	Thr	Tyr	Ser	Arg	Tyr	Ala	Asp	Asp	Ile	Thr	Ile	Ser	Thr
				165					170					175	
Asn	Lys	Asn	Thr	Phe	Pro	Leu	Glu	Met	Ala	Thr	Val	Gln	Pro	Glu	Gly
			180					185					190		
Val	Val	Leu	Gly	Lys	Val	Leu	Val	Lys	Glu	Ile	Glu	Asn	Ser	Gly	Phe
		195					200					205			
Glu	Ile	Asn	Asp	Ser	Lys	Thr	Arg	Leu	Thr	Tyr	Lys	Thr	Ser	Arg	Gln
	210					215					220				
Glu	Val	Thr	Gly	Leu	Thr	Val	Asn	Arg	Ile	Val	Asn	Ile	Asp	Arg	Cys
225					230					235					240
Tyr	Tyr	Lys	Lys	Thr	Arg	Ala	Leu	Ala	His	Ala	Leu	Tyr	Arg	Thr	Gly
				245					250					255	
Glu	Tyr	Lys	Val	Pro	Asp	Glu	Asn	Gly	Val	Leu	Val	Ser	Gly	Gly	Leu
			260					265					270		
Asp	Lys	Leu	Glu	Gly	Met	Phe	Gly	Phe	Ile	Asp	Gln	Val	Asp	Lys	Phe
		275					280					285			
Asn	Asn	Ile	Lys	Lys	Lys	Leu	Asn	Lys	Gln	Pro	Asp	Arg	Tyr	Val	Leu
	290					295					300				
Thr	Asn	Ala	Thr	Leu	His	Gly	Phe	Lys	Leu	Lys	Leu	Asn	Ala	Arg	Glu
305					310					315					320

09-0386

Lys Ala Tyr Ser Lys Phe Ile Tyr Tyr Lys Phe Phe His Gly Asn Thr
 325 330 335
 Cys Pro Thr Ile Ile Thr Glu Gly Lys Thr Asp Arg Ile Tyr Leu Lys
 340 345 350
 Ala Ala Leu His Ser Leu Glu Thr Ser Tyr Pro Glu Leu Phe Arg Glu
 355 360 365
 Lys Thr Asp Ser Lys Lys Lys Glu Ile Asn Leu Asn Ile Phe Lys Ser
 370 375 380
 Asn Glu Lys Thr Lys Tyr Phe Leu Asp Leu Ser Gly Gly Thr Ala Asp
 385 390 395 400
 Leu Lys Lys Phe Val Glu Arg Tyr Lys Asn Asn Tyr Ala Ser Tyr Tyr
 405 410 415
 Gly Ser Val Pro Lys Gln Pro Val Ile Met Val Leu Asp Asn Asp Thr
 420 425 430
 Gly Pro Ser Asp Leu Leu Asn Phe Leu Arg Asn Lys Val Lys Ser Cys
 435 440 445
 Pro Asp Asp Val Thr Glu Met Arg Lys Met Lys Tyr Ile His Val Phe
 450 455 460
 Tyr Asn Leu Tyr Ile Val Leu Thr Pro Leu Ser Pro Ser Gly Glu Gln
 465 470 475 480
 Thr Ser Met Glu Asp Leu Phe Pro Lys Asp Ile Leu Asp Ile Lys Ile
 485 490 495
 Asp Gly Lys Lys Phe Asn Lys Asn Asn Asp Gly Asp Ser Lys Thr Glu
 500 505 510
 Tyr Gly Lys His Ile Phe Ser Met Arg Val Val Arg Asp Lys Lys Arg
 515 520 525
 Lys Ile Asp Phe Lys Ala Phe Cys Cys Ile Phe Asp Ala Ile Lys Asp
 530 535 540
 Ile Lys Glu His Tyr Lys Leu Met Leu Asn Ser
 545 550 555

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 243 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

030397

~~CONFIDENTIAL~~

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 241 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Arg Trp Phe Ala Phe His Arg Glu Val Asp Thr Ala Thr His Tyr Val
1 5 10 15
Ser Trp Thr Ile Pro Lys Arg Asp Gly Ser Lys Arg Thr Ile Thr Ser
20 25 30
Pro Lys Pro Glu Leu Lys Ala Ala Gln Arg Trp Val Leu Ser Asn Val
35 40 45
Val Glu Arg Leu Pro Val His Gly Ala Ala His Gly Phe Val Ala Gly
50 55 60
Arg Ser Ile Leu Thr Asn Ala Leu Ala His Gln Gly Ala Asp Val Val
65 70 75 80
Val Lys Val Asp Leu Lys Asp Phe Phe Pro Ser Val Thr Trp Arg Arg
85 90 95
Val Lys Gly Leu Leu Arg Lys Gly Gly Leu Arg Glu Gly Thr Ser Thr
100 105 110
Leu Leu Ser Leu Leu Ser Thr Glu Ala Pro Arg Glu Ala Val Gln Phe
115 120 125
Pro Arg Glu Leu Leu His Val Ala Lys Gly Pro Arg Ala Leu Pro Gln
130 135 140
Gly Ala Pro Thr Ser Pro Gly Ile Thr Asn Ala Leu Cys Leu Lys Leu
145 150 155 160
Asp Lys Arg Leu Ser Ala Leu Ala Lys Arg Leu Gly Phe Thr Tyr Thr
165 170 175
Arg Tyr Ala Asp Asp Leu Thr Phe Ser Trp Thr Lys Ala Lys Gln Pro
180 185 190
Lys Pro Arg Arg Thr Gln Arg Pro Pro Val Ala Val Leu Leu Ser Arg
195 200 205
Val Gln Glu Val Val Glu Ala Glu Gly Phe Arg Val His Pro Asp Lys
210 215 220
Thr Arg Val Ala Arg Lys Gly Thr Arg Gln Arg Val Thr Gly Leu Val
225 230 235 240

00808031.030397

Val

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 231 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Arg His Tyr Ser Ile His Arg Pro Arg Glu Arg Val Arg His Tyr Val
1 5 10 15
Thr Phe Ala Val Pro Lys Arg Ser Gly Gly Val Arg Leu Leu His Ala
20 25 30
Pro Lys Arg Arg Leu Lys Ala Leu Gln Arg Arg Met Leu Ala Leu Leu
35 40 45
Val Ser Lys Leu Pro Val Ser Pro Gln Ala His Gly Phe Val Pro Gly
50 55 60
Arg Ser Ile Lys Thr Gly Ala Ala Pro His Val Gly Arg Arg Val Val
65 70 75 80
Leu Lys Leu Asp Leu Lys Asp Phe Phe Pro Ser Val Thr Phe Ala Arg
85 90 95
Val Arg Gly Leu Leu Lys Ala Leu Gly Tyr Gly Tyr Pro Val Ala Ala
100 105 110
Thr Leu Ala Val Leu Met Thr Glu Ser Glu Arg Gln Pro Val Glu Leu
115 120 125
Glu Gly Ile Leu Phe His Val Pro Val Gly Pro Arg Val Cys Val Gln
130 135 140
Gly Ala Pro Thr Ser Pro Ala Leu Cys Asn Ala Val Leu Leu Arg Leu
145 150 155 160
Asp Arg Arg Leu Ala Gly Leu Ala Arg Arg Tyr Gly Tyr Thr Tyr Thr
165 170 175
Arg Tyr Ala Asp Asp Leu Thr Phe Ser Gly Asp Asp Val Thr Ala Leu
180 185 190
Glu Arg Val Arg Ala Leu Ala Ala Arg Tyr Val Gln Glu Glu Gly Phe
195 200 205

0808031.030397

Glu Val Asn Arg Glu Lys Thr Arg Val Gln Arg Arg Gly Gly Ala Gln
 210 215 220

Arg Val Thr Gly Val Thr Val
 225 230

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 234 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Phe Leu Thr Asn Val Leu Tyr Arg Ile Gly Ser Asp Asn Gln Tyr Thr
 1 5 10 15
 Gln Phe Thr Ile Pro Lys Lys Gly Lys Gly Val Arg Thr Ile Ser Ala
 20 25 30
 Pro Thr Asp Arg Leu Lys Asp Ile Gln Arg Arg Ile Cys Asp Leu Leu
 35 40 45
 Ser Asp Cys Arg Asp Glu Ile Phe Ala Ile Arg Lys Ile Ser Asn Asn
 50 55 60
 Tyr Ser Phe Gly Phe Glu Arg Gly Lys Ser Ile Ile Leu Asn Ala Tyr
 65 70 75 80
 Lys His Arg Gly Lys Gln Ile Ile Leu Asn Ile Asp Leu Lys Asp Phe
 85 90 95
 Phe Glu Ser Phe Asn Phe Gly Arg Val Arg Gly Tyr Phe Leu Ser Asn
 100 105 110
 Gln Asp Phe Leu Leu Asn Pro Val Val Ala Thr Thr Leu Ala Lys Ala
 115 120 125
 Ala Cys Tyr Asn Gly Thr Leu Pro Gln Gly Ser Pro Cys Ser Pro Ile
 130 135 140
 Ile Ser Asn Leu Ile Cys Asn Ile Met Asp Met Arg Leu Ala Lys Leu
 145 150 155 160
 Ala Lys Lys Tyr Gly Cys Thr Tyr Ser Arg Tyr Ala Asp Asp Ile Thr
 165 170 175
 Ile Ser Thr Asn Lys Asn Thr Phe Pro Leu Glu Met Ala Thr Val Gln
 180 185 190

26E0E0-TE080840

Pro Glu Gly Val Val Leu Gly Lys Val Leu Val Lys Glu Ile Glu Asn
 195 200 205

Ser Gly Phe Glu Ile Asn Asp Ser Lys Thr Arg Leu Thr Tyr Lys Thr
 210 215 220

Ser Arg Gln Glu Val Thr Gly Leu Thr Val
 225 230

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 215 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Val Glu Thr Leu Arg Leu Leu Ile Tyr Thr Ala Asp Phe Arg Tyr Arg
 1 5 10 15

Ile Tyr Thr Val Glu Lys Lys Gly Pro Glu Lys Arg Met Arg Thr Ile
 20 25 30

Tyr Gln Pro Ser Arg Glu Leu Lys Ala Leu Gln Gly Trp Val Leu Arg
 35 40 45

Asn Ile Leu Asp Lys Leu Ser Ser Ser Pro Phe Ser Ile Gly Phe Glu
 50 55 60

Lys His Gln Ser Ile Leu Asn Asn Ala Thr Pro His Ile Gly Ala Asn
 65 70 75 80

Phe Ile Leu Asn Ile Asp Leu Glu Asp Phe Phe Pro Ser Leu Thr Ala
 85 90 95

Asn Lys Val Phe Gly Val Phe His Ser Leu Gly Tyr Asn Arg Leu Ile
 100 105 110

Ser Ser Val Leu Thr Lys Ile Cys Cys Tyr Lys Asn Leu Leu Pro Gln
 115 120 125

Gly Ala Pro Ser Ser Pro Lys Leu Ala Asn Leu Ile Cys Ser Lys Leu
 130 135 140

Asp Tyr Arg Ile Gln Gly Tyr Ala Gly Ser Arg Gly Leu Ile Tyr Thr
 145 150 155 160

Arg Tyr Ala Asp Asp Leu Thr Leu Ser Ala Gln Ser Met Lys Lys Val
 165 170 175

26E0E0-TE080880

Asp Ile Ile Val Ser Thr Asn Met Lys Gly Ala Ser Lys Leu Ile Leu
180 185 190

Asp Cys Phe Lys Arg Thr Met Lys Glu Ile Gly Pro Asp Phe Lys Ile
195 200 205

Asn Ile Lys Lys Phe Lys Ile Cys Ser Ala Ser Gly Gly Ser Ile Val
210 215 220

Val Thr Gly Leu Lys Val
225 230

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Ile Gln Arg Leu His Ala Leu Ser Asn His Ala Gly Arg His Tyr Arg
1 5 10 15

Arg Ile Ile Leu Ser Lys Arg His Gly Gly Gln Arg Leu Val Leu Ala
20 25 30

Pro Asp Tyr Leu Leu Lys Thr Val Gln Arg Asn Ile Leu Lys Asn Val
35 40 45

Leu Ser Gln Phe Pro Leu Ser Pro Phe Ala Thr Ala Tyr Arg Pro Gly
50 55 60

Cys Pro Ile Val Ser Asn Ala Gln Pro His Cys Gln Gln Pro Gln Ile
65 70 75 80

Leu Lys Leu Asp Ile Glu Asn Phe Phe Asp Ser Ile Ser Trp Leu Gln
85 90 95

Val Trp Arg Val Phe Arg Gln Ala Gln Leu Pro Arg Asn Val Val Thr
100 105 110

Met Leu Thr Trp Ile Cys Cys Tyr Asn Asp Ala Leu Pro Gln Gly Ala
115 120 125

Pro Thr Ser Pro Ala Ile Ser Asn Leu Val Met Arg Arg Phe Asp Glu
130 135 140

Arg Ile Gly Glu Trp Cys Gln Ala Arg Gly Ile Thr Tyr Thr Arg Tyr
145 150 155 160

030397

55

60

65

CGC Arg 70	CGC Arg	TAC Tyr	ACC Thr	CCG Pro	GGC Gly 75	CGG Arg	AAG Lys	AAG Lys	TGG Trp	ATG Met 80	GAG Glu	GCC Ala	GCC Ala	GAG Glu	GCC Ala 85	533
CGG Arg	CGG Arg	CTG Leu	TTC Phe	TCC Ser 90	GCC Ala	ACG Thr	CTG Leu	CGC Arg	ACG Thr 95	CGG Arg	AAC Asn	CGG Arg	AAC Asn	CTG Leu 100	AGG Arg	581
GAC Asp	TTG Leu	CTG Leu	CCC Pro 105	GAC Asp	GAG Glu	GCA Ala	CAG Gln	CTG Leu 110	GCG Ala	CGC Arg	TAC Tyr	GGC Gly	CTG Leu 115	CCG Pro	GTC Val	629
TGG Trp	CGC Arg	ACG Thr 120	GAA Glu	GAG Glu	GAC Asp	GTG Val	GCA Ala	GCG Ala	GCC Ala	CTG Leu	GGC Gly	GTC Val 130	TCG Ser	GTG Val	GGC Gly	677
GTG Val 135	CTC Leu	CGC Arg	CAC His	TAC Tyr	AGC Ser	ATC Ile 140	CAC His	CGC Arg	CCG Pro	CGC Arg	GAG Glu 145	CGG Arg	GTG Val	CGG Arg	CAC His	725
TAC Tyr 150	GTG Val	ACC Thr	TTC Phe	GCC Ala	GTG Val 155	CCC Pro	AAG Lys	CGC Arg	TCC Ser	GGA Gly 160	GGC Gly	GTC Val	CGG Arg	CTG Leu	CTG Leu 165	773
CAT His	GCG Ala	CCC Pro	AAG Lys	CGG Arg 170	CGC Arg	CTG Leu	AAG Lys	GCC Ala	CTG Leu 175	CAA Gln	CGC Arg	CGG Arg	ATG Met	CTG Leu 180	GCG Ala	821
CTC Leu	CTG Leu	GTG Val	TCG Ser 185	AAG Lys	CTC Leu	CCC Pro	GTG Val	AGT Ser 190	CCA Pro	CAG Gln	GCC Ala	CAT His	GGC Gly 195	TTC Phe	GTG Val	869
CCC Pro 200	GGC Gly	CGC Arg	TCC Ser	ATC Ile	AAG Lys	ACG Thr	GGC Gly 205	GCC Ala	GCG Ala	CCG Pro	CAC His	GTG Val 210	GGC Gly	CGG Arg	CGG Arg	917
GTG Val 215	GTC Val	CTG Leu	AAG Lys	CTG Leu	GAC Asp	CTG Leu 220	AAG Lys	GAC Asp	TTC Phe	TTC Phe	CCC Pro 225	TCC Ser	GTC Val	ACC Thr	TTC Phe	965
GCG Ala 230	CGG Arg	GTG Val	CGA Arg	GGG Gly	CTG Leu 235	CTC Leu	ATC Ile	GCC Ala	CTG Leu	GGC Gly 240	TAC Tyr	GGC Gly	TAT Tyr	CCC Pro	GTG Val 245	1013
GCG Ala	GCC Ala	ACG Thr	CTC Leu	GCG Ala 250	GTG Val	CTG Leu	ATG Met	ACG Thr	GAG Glu 255	TCC Ser	GAG Glu	CGC Arg	CAG Gln	CCC Pro 260	GTG Val	1061
GAG Glu	CTG Leu	GAG Glu	GGC Gly 265	ATC Ile	CTC Leu	TTC Phe	CAC His	GTT Val 270	CCC Pro	GTG Val	GGC Gly	CCA Pro	CGC Arg 275	GTC Val	TGC Cys	1109

03308031-03009
460000-TE008000

0303037 0303037 0303037 0303037

(i) SEQUENCE CHARACTERISTICS:

- (ix) FEATURE:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CTCCGAGCCC	GCCTCCGAGG	ACGCGCTCGC	GGCCCCGGGCG	GCGGGGGCGG	ACGCGCGGCG												60
GCGGCCCACG	GAGACGCTTG	ACCCGGGAGA	CGACGAATGA	CGATAACGGC	AGGTGCTCTC												120
GGGAGAGGCC	AGGGCTCGCA	GATGAGCCAT	GAGTACCGCG	GTGTTTCGCC	GCGGGGGTGT												180
TCTGTCCCCA	TCTCTTCGCC	AGGGTCCCAG	CGTACGCAAC	GCAGGGAGCC	CCGGGTCCAA												240
CGCCTCGCAG	GTCGTCCCCT	GGCCTCTTCC	GGAGCACC	ATG	AGC	TGG	TTC	GAC									293
					Met	Ser	Trp	Phe	Asp								
					1				5								
ACC	ACC	CTC	TCC	CGG	CTC	AAG	GGG	TTG	TTC	AGC	CGT	CCC	GTG	ACA	CGA		341
Thr	Thr	Leu	Ser	Arg	Leu	Lys	Gly	Leu	Phe	Ser	Arg	Pro	Val	Thr	Arg		
				10					15					20			
AGC	ACC	ACC	GGG	CTG	GAC	GTG	CCG	CTG	GAT	GCC	CAC	GGA	CGT	CCC	CAG		389
Ser	Thr	Thr	Gly	Leu	Asp	Val	Pro	Leu	Asp	Ala	His	Gly	Arg	Pro	Gln		
			25					30					35				
GAC	GTC	GTG	ACG	GAG	ACG	GTC	TCC	ACG	TCG	GGC	CCC	CTG	AAG	CCA	GGG		437
Asp	Val	Val	Thr	Glu	Thr	Val	Ser	Thr	Ser	Gly	Pro	Leu	Lys	Pro	Gly		
		40					45					50					
CAC	CTG	CGA	CAG	GTC	CGC	CGG	GAT	GCG	CGG	CTG	CTC	CCC	AAG	GGC	GTC		485
His	Leu	Arg	Gln	Val	Arg	Arg	Asp	Ala	Arg	Leu	Leu	Pro	Lys	Gly	Val		

GTG CAG GGC GCC CCC ACG AGC CCC GCC CTG TGC AAC GCG GTG CTG CTG 1157
Val Gln Gly Ala Pro Thr Ser Pro Ala Leu Cys Asn Ala Val Leu Leu
280 285 290

CGA CTG GAC CGG CGG CTG GCG GGA CTG GCG CGT CGG TAC GGC TAC ACG 1205
Arg Leu Asp Arg Arg Leu Ala Gly Leu Ala Arg Arg Tyr Gly Tyr Thr
295 300 305

TAC ACG CGC TAC GCG GAT GAC CTC ACC TTC TCC GGC GAC GAC GTC ACG 1253
Tyr Thr Arg Tyr Ala Asp Asp Leu Thr Phe Ser Gly Asp Asp Val Thr
310 315 320 325

GCG CTG GAG CGA GTC CGC GCG CTG GCC GCG CGG TAC GTG CAG GAG GAA 1301
Ala Leu Glu Arg Val Arg Ala Leu Ala Ala Arg Tyr Val Gln Glu Glu
330 335 340

GGC TTC GAG GTC AAC CGC GAG AAG ACC CGC GTG CAG CGC CGG GGC GGT 1349
Gly Phe Glu Val Asn Arg Glu Lys Thr Arg Val Gln Arg Arg Gly Gly
345 350 355

GCC CAG CGC GTC ACT GGC GTC ACC GTG AAT ACG ACG CTG GGC TTG TCA 1397
Ala Gln Arg Val Thr Gly Val Thr Val Asn Thr Thr Leu Gly Leu Ser
360 365 370

CGC GAG GAG CGG CCG CGG CTC CGG GCG ATG CTG CAC CAG GAG GCG CGG 1445
Arg Glu Glu Arg Pro Arg Leu Arg Ala Met Leu His Gln Glu Ala Arg
375 380 385

TCG GAG GAC GTC GAG GCA CAC CGC GCG CAC CTC GAC GGC CTC CTG GCC 1493
Ser Glu Asp Val Glu Ala His Arg Ala His Leu Asp Gly Leu Leu Ala
390 395 400 405

TAC GTG AAG ATG CTC AAC CCG GAG CAG GCG GAG CGG CTC GCT CGC CGG 1541
Tyr Val Lys Met Leu Asn Pro Glu Gln Ala Glu Arg Leu Ala Arg Arg
410 415 420

CGC AAG CCG CGC GGG ACG TGAGCGAGGG CTCAGCTCCG GATGGGCCAG 1589
Arg Lys Pro Arg Gly Thr
425

GGCCTGTCAC GCGTCCCGGC CTCCCAGTTG TCATGGCGGC CGTCCCAGTA C 1640

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3060 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 763..2202

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CCCACTTCCG GCGCTCGGGC TGC GCGAGGG CCCGTGCGAG CACATGATGG CGCTGCGGCT 60
CGTCCAGGTC CGGCACCGCG CCGAGCAGGA AGCACTGCGT CAGACCCCCG CGGGCCGCCA 120
GCTCATCCGC GCGGAGACGC GCTCCTACGT GCGGCGCGAG CCCTCCGGCC AGGAGCAGGT 180
GTACCGCGTC TCATTGGATG GGAAAGTGGT GGCGGTGGAG TGGGGCCCCC GCCAGGGGGA 240
GTCCCGCCGG CAGAAGCTCT GGTTCGACAC GGACGCCGAG GCGCGCACCG CCTACTTCAC 300
GCGCCTGGAG TCCTTGGCCG CGGAGGGATA TATCGATGCG GCTGCTTCAA TGATGTAGAA 360
CACGCAAGCC ACGGGGCCGC GGGCGCGCGG CGGAAAGGCA GGTGCGACGG AACGACAGAC 420
ACTCGTGCGA GCGACCGAGA GAGGTCCCAA GCCATCAGCC TCAGCGCCTC GAGCGCGAGA 480
GCGGCGTTGC GCCGCTCTGG TTGAATTGCA GGACACTCTC CGCAAGGTAG CCTGTTCTTG 540
GCTCTCTTCC CTCCGGTGAG TACCTCTCCG GCCGGGGAGC TGAACCAACG ACGCAACCGC 600
CGTTTCCCCG GCCGGAGAGG TACTCACCGG AGGGGAGAGC CGGTGAGGCT ACCGTGCCCC 660
AGGTGAGAAG GTGGTGCCTT CGGGCCTCCC TCGACCGCTC GCGCTCCGTC GCCCTGCCCT 720
GCCTCGCCCC CCCCACCTTG CTCACCGGCG CCAGGAGCCG TC ATG ACC GCC AAG 774
Met Thr Ala Lys
1
CTG GAG TCA CAC GTC CCC GCC GCG CCC CCC GTC TCC GCC GAG GCG CCC 822
Leu Glu Ser His Val Pro Ala Ala Pro Pro Val Ser Ala Glu Ala Pro
5 10 15 20
GCC CCC ACC CGT CCC GAT GCC GCG AAG CAG GAG GCC CGC CGC GCC CAC 870
Ala Pro Thr Arg Pro Asp Ala Ala Lys Gln Glu Ala Arg Arg Ala His
25 30 35
CAC GAG GCG CTG CGC CTG CGG TGG AAG GCC ATC GAA GAG GCG GGC GGC 918
His Glu Ala Leu Arg Leu Arg Trp Lys Ala Ile Glu Glu Ala Gly Gly
40 45 50
ACG GAC GCC TGG GTG CGG CAG CAG CTG GTG GCC AAG GGC GTC GCG GCG 966
Thr Asp Ala Trp Val Arg Gln Gln Leu Val Ala Lys Gly Val Ala Ala
55 60 65
GAA GAG GTG GAC TTC GAG TCG CTC AGC GAC AAG CAG AAG GCG GCC TGG 1014
Glu Glu Val Asp Phe Glu Ser Leu Ser Asp Lys Gln Lys Ala Ala Trp
70 75 80
AAG GAG AAG AAG AAG GCC GAG GCC ACC GAG CGG CGC GCG CAG AAG CGC 1062

GGG Gly	GCC Ala 310	CCC Pro	ACC Thr	TCT Ser	CCG Pro	GCG Ala 315	CTG Leu	ACG Thr	AAC Asn	GCG Ala	CTG Leu 320	TGC Cys	CTG Leu	CGG Arg	CTG Leu	1734	
GAC Asp 325	AAG Lys	CGG Arg	CTC Leu	TCG Ser	GCG Ala 330	CTG Leu	TCG Ser	AAG Lys	CGG Arg	CTG Leu 335	GGC Gly	TTC Phe	ACG Thr	TAC Tyr	ACG Thr 340	1782	
CGC Arg	TAT Tyr	GCG Ala	GAT Asp	GAC Asp 345	CTG Leu	ACG Thr	TTC Phe	TCC Ser	TGG Trp 350	CGG Arg	CGG Arg	GCG Ala	AAG Lys	AAG Lys 355	TCC Ser	1830	
CGG Arg	CAG Gln	AAG Lys	GAA Glu 360	CTC Leu	CCC Pro	CTG Leu	GCG Ala	GAT Asp 365	GCG Ala	CCG Pro	GTG Val	GCG Ala	CTG Leu 370	CTC Leu	CTG Leu	1878	
GCG Ala	CGG Arg	GTG Val 375	AAG Lys	GGT Gly	GTG Val	CTG Leu	GAG Glu 380	GCC Ala	GAG Glu	GGT Gly	TTC Phe	ACG Thr 385	CTG Leu	CAC His	CCG Pro	1926	
GAC Asp	AAG Lys 390	ACG Thr	CGG Arg	GTG Val	CAG Gln	CGC Arg 395	AAG Lys	GGC Gly	AGC Ser	CGG Arg	CAG Gln 400	CGG Arg	GTG Val	ACG Thr	GGG Gly	1974	
CTC Leu 405	GTG Val	GTG Val	AAC Asn	GAG Glu	GCC Ala 410	CCC Pro	GAG Glu	GGC Gly	GTT Val	CCG Pro 415	GGT Gly	GCC Ala	CGG Arg	GTG Val	CCC Pro 420	2022	
CGC Arg	GAT Asp	GTG Val	GTG Val	CGG Arg 425	CGG Arg	CTG Leu	CGC Arg	GCG Ala	GCG Ala 430	ATC Ile	CAC His	AAC Asn	CGG Arg	GAG Glu 435	CAG Gln	2070	
GGC Gly	AAG Lys	CCC Pro	GGC Gly 440	CCC Pro	ACC Thr	GGG Gly	GAG Glu	ACG Thr 445	CTG Leu	GAG Glu	CAG Gln	CTC Leu	AAG Lys 450	GGG Gly	CTC Leu	2118	
GCG Ala	GCC Ala	TTC Phe 455	CTT Leu	CAC His	ATG Met	ACG Thr	GAC Asp 460	GCG Ala	GAG Glu	AAG Lys	GGC Gly	CGC Arg 465	GCC Ala	TTC Phe	CTG Leu	2166	
CGA Arg	CGG Arg 470	CTG Leu	GAG Glu	GCC Ala	CTC Leu	GAG Glu 475	AAG Lys	CGC Arg	CAG Gln	ACC Thr	GCC Ala 480	TGACCCTCAC				2212	
TGGTCGTCCG		GGGCATCGCA				GCGGGCGCCG				GGACGGACCG				TCACCCCCCA		GATCTCCATG	2272
CCATGCTGGG		GATTCTGGGC				GGTGAAGAAG				ACTTCCCAGC				CGAGACGGAC		GAAGCCCTGC	2332
GGATCCGATG		ACTCCTCGCC				CGGGGCGATC				TCCCGGAGGG				GCACCGTTCC		GACGTCCGTG	2392
CCATTGCTCA		CCCAGGGCTC				CCGGCCCCAG				CCTTGGGTGT				CCGCCGAGAA		GAAGAGCAGC	2452
CCGGAGATGG		CCGTCAGGTT				CTCCGGCGAC				GCATCCTCGG				GGCCCGGCGC		CAAATCCTTC	2512

AGCAGCAGGG TGCCCTTGGC GGTGCCATCG CTGGACCACA GCTCCCGGCC GTGGAGGCTG 2572
TCACTCGCGG CGAAGTAGAG CATCCCATTC AGCGCCTTGA TGGCGCTGGG CGCCGAGCTG 2632
TCCGGACCCG GCCAGATGTC CTTACCCCGG ACCGTGCCAT GCGACGTGCC ATCGCTGACC 2692
CACAGCTCCT CGCCCTCGGG CTGGCCCCAG AACTCGGGCT CGCCTCCCCC GGCGCTGAAG 2752
AAGATCTTCC CCCCAGAGCGC CGTGAGATCA TGC GGATAGA GGCCGGGGAA GAAGCGCAGC 2812
TGCTCGGAGA CGGTGCCTCT GGAGCACCAC AGGCTGGCCT CGCCTTCGTC ATTGTCGAGC 2872
AGGAAGAAGA GCACCGAGTC CGCCGCGGTG AACGCGGAGA GGAAGTTGTC CTCGGGGCCC 2932
GTGAAGACAG ACGTGGTGCT GGACAGCCCC AGGCTGCGCC AGATGAACAC CTCGTCATTG 2992
ACGTTGGCCA CGAAGAAGAG CGCATCGCCG ACCCGGGTGA GCCGGCGCGG GCTGGAGCTG 3052
CCGGGCAC 3060

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2788 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..103

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 707..1654

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1644..2591

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

T TTC GAG AAG CGC CAT ACC AAA CAG GGG ATA CAG ACC AAC CTG ACG 46
Phe Glu Lys Arg His Thr Lys Gln Gly Ile Gln Thr Asn Leu Thr
1 5 10 15
CTG AAA GAG GAA AGC TAC GGC GAC TGG CTG CCG AAG TGC GAC GAC CCC 94
Leu Lys Glu Glu Ser Tyr Gly Asp Trp Leu Pro Lys Cys Asp Asp Pro
20 25 30
GCA GCA ACA TAACCTCACT CAGACCGGCA ACAGCCGGTC TTTTCCTTTC 143

2572-3060

Ala Ala Thr

TGGCCATTGC CACAAGGTGA ACAATCCACT GTTCACCCTT CACCGTTTAT TCACCCTTTA 203
TCACTATGAA ATTATTAATA AAAAACCAGA GGTGAACAGT GTGAACAGTA AAACCTGAAA 263
AAACTTTTTTA TCACCCCGCG CATCGCCCGA CTGGACAGAT CCAGAACGAG CAAAAATCAC 323
AAAGGTGACG AGTCGACTGT TCACTCTTCA CCAACTCATC ACCACCTAAC CACATGATAT 383
AAAATGATAA ATAATCGAGG TGAACAGTTA AATGCAAAAA AACTTTTTTCT CAGCTCTTGG 443
ATAAAAGAAA ATTAATTCAC ATCAATAGCT TTCCTCTTGA ATCCTCTTGA GGTTTATGAG 503
AGCGTAACAG AGCCAAACCT AGCATTTTAT GGGTTAATAG CCCATCGCGC ATGAGTCATG 563
GTTTCGCCTA GTATTTTAGC TATGCCCGTC GTTCAGTTCG CTGAGCGGCG GCTGGGGGCC 623
ACCGATCAGC GAACTGATCG ACGTGCTCAA GTAGGTTTGG CTCTTTTAGT CCTCTACCAT 683
CAAGGTGCAT AAGGATATTC TCG ATG CTG ACT CAG CTA AAA AAA AAT GGT 733
Met Leu Thr Gln Leu Lys Lys Asn Gly
1 5
ACT GAG GTA TCT AGA GCA ACC GCG TTA TTT TCA TCA TTC GTT GAA AAG 781
Thr Glu Val Ser Arg Ala Thr Ala Leu Phe Ser Ser Phe Val Glu Lys
10 15 20 25
AAC AAA GTA AAA TGT CCT GGT AAT GTA AAA AAA TTC GTC TTT CTG TGT 829
Asn Lys Val Lys Cys Pro Gly Asn Val Lys Lys Phe Val Phe Leu Cys
30 35 40
GGT GCT AAC AAA AAC AAT GGA GAA CCA TCA GCA AGA CGA TTG GAA TTA 877
Gly Ala Asn Lys Asn Asn Gly Glu Pro Ser Ala Arg Arg Leu Glu Leu
45 50 55
ATA AAT TTT TCT GAA AGG TAT TTG AAT AAC TGT CAC TTT TTT CTT GCT 925
Ile Asn Phe Ser Glu Arg Tyr Leu Asn Asn Cys His Phe Phe Leu Ala
60 65 70
GAA CTA GTT TTC AAA GAA TTA AGC ACC GAT GAA GAA TCA TTA TCT GAT 973
Glu Leu Val Phe Lys Glu Leu Ser Thr Asp Glu Glu Ser Leu Ser Asp
75 80 85
AAT TTA TTA GAT ATC GAA GCT GAC TTA TCT AAA TTA GCT GAT CAT ATT 1021
Asn Leu Leu Asp Ile Glu Ala Asp Leu Ser Lys Leu Ala Asp His Ile
90 95 100 105
ATC ATT GTT TTA GAA AGT TAT TCA TCT TTC ACG GAA CTT GGT GCA TTC 1069
Ile Ile Val Leu Glu Ser Tyr Ser Ser Phe Thr Glu Leu Gly Ala Phe
110 115 120
GCA TAC AGC AAG CAA TTA CGC AAG AAA TTA ATA ATA GTT AAC AAT ACA 1117

Ala	Tyr	Ser	Lys	Gln	Leu	Arg	Lys	Lys	Leu	Ile	Ile	Val	Asn	Thr		
			125					130					135			
AAA	TTT	ATA	AAT	GAG	AAA	TCA	TTT	ATA	AAT	ATG	GGA	CCA	ATA	AAG	GCT	1165
Lys	Phe	Ile	Asn	Glu	Lys	Ser	Phe	Ile	Asn	Met	Gly	Pro	Ile	Lys	Ala	
		140					145					150				
ATT	ACT	CAG	CAA	TCA	CAA	CAA	TCT	GGT	CAT	TTC	TTA	CAT	TAT	AAA	ATG	1213
Ile	Thr	Gln	Gln	Ser	Gln	Gln	Ser	Gly	His	Phe	Leu	His	Tyr	Lys	Met	
		155					160				165					
ACA	GAA	GGT	ATT	GAA	AGT	ATA	GAG	CGC	TCT	GAT	GGG	ATT	GGC	GAA	ATA	1261
Thr	Glu	Gly	Ile	Glu	Ser	Ile	Glu	Arg	Ser	Asp	Gly	Ile	Gly	Glu	Ile	
		170			175					180					185	
TTC	GAC	CCC	CTA	TAT	GAT	ATT	CTT	TCT	AAG	AAC	GAC	AGA	GCA	ATT	TCA	1309
Phe	Asp	Pro	Leu	Tyr	Asp	Ile	Leu	Ser	Lys	Asn	Asp	Arg	Ala	Ile	Ser	
			190						195					200		
AGA	ACT	TTA	AAA	AAA	GAA	GAG	TTA	GAT	CCT	TCC	AGT	AAC	TTC	AAT	AAA	1357
Arg	Thr	Leu	Lys	Lys	Glu	Glu	Leu	Asp	Pro	Ser	Ser	Asn	Phe	Asn	Lys	
			205					210					215			
GAC	TCA	GTA	CGA	TTT	ATT	CAT	GAC	GTA	ATT	TTT	GTA	TGT	GGT	CCT	TTG	1405
Asp	Ser	Val	Arg	Phe	Ile	His	Asp	Val	Ile	Phe	Val	Cys	Gly	Pro	Leu	
		220					225					230				
CAA	CTT	AAT	GAA	CTC	ATC	GAA	ATA	ATC	ACA	AAA	ATA	TTT	GGC	ACA	GAA	1453
Gln	Leu	Asn	Glu	Leu	Ile	Glu	Ile	Ile	Thr	Lys	Ile	Phe	Gly	Thr	Glu	
		235				240					245					
AGC	CAT	TAC	AAA	AAA	AAT	CTT	CTA	AAG	CAC	CTT	GGT	ATT	CTA	ATA	GCT	1501
Ser	His	Tyr	Lys	Lys	Asn	Leu	Leu	Lys	His	Leu	Gly	Ile	Leu	Ile	Ala	
					255					260					265	
ATT	AGA	ATA	ATA	TCA	TGC	ACA	AAT	GGG	ATT	TAT	TAT	TCT	TTG	TAT	AAA	1549
Ile	Arg	Ile	Ile	Ser	Cys	Thr	Asn	Gly	Ile	Tyr	Tyr	Ser	Leu	Tyr	Lys	
				270				275						280		
GAA	TAT	TAT	TTT	AAA	TAT	GAC	TTT	GAC	ATT	GAC	AAC	ATA	TCA	TCA	ATG	1597
Glu	Tyr	Tyr	Phe	Lys	Tyr	Asp	Phe	Asp	Ile	Asp	Asn	Ile	Ser	Ser	Met	
			285					290					295			
TTT	AAA	GTT	TTT	TTC	CTC	AAG	AAC	AAG	CCA	GAA	AGG	ATG	AGG	GTA	TAT	1645
Phe	Lys	Val	Phe	Phe	Leu	Lys	Asn	Lys	Pro	Glu	Arg	Met	Arg	Val	Tyr	
		300					305					310				
GAG	AAT	ATA	TAGCCTAATT			GATTCTCAGA			CATTGATGAC			TAAGGGATTT				1694
Glu	Asn	Ile														
		315														
GCTTCTGAAG			TAATGCGATC			ACCTGAGCCG			CCAAAAAAT</							

ATGAATAATG TTTTTTCGAA GCTCCCAATG CATAATGCTG CATATGCATT TGTAAAAAAC 1874
CGATCAATAA AAAGCAATGC TTTATTACAT GCCGAATCAA AGAATAAGTA TTATGTGAAA 1934
ATAGATCTCA AAGATTTTTT CCCTTCAATA AAATTTACTG ATTTTGAGTA CGCATTCACT 1994
CGTTATCGAG ATCGCATTGA ATTTACTACA GAATATGATA AGGAGTTACT ACAACTTATA 2054
AAAACGATCT GCTTTATATC AGATAGCACT CTCCCTATCG GGTTCCTAC ATCTCCATTA 2114
ATTGCAAACCT TTGTGGCAAG AGAACTTGAT GAAAAACTGA CGCAAAACCT AAATGCAATT 2174
GATAAACTTA ATGCCACTTA TACACGATAT GCTGATGATA TTATTGTCTC TACAAATATG 2234
AAAGGGGCTA GCAAATTAAT TCTGGATTGT TTTAAAAGAA CAATGAAAGA GATTGGTCCA 2294
GACTTTAAAA TTAACATTAA AAAATTTAAG ATTTGTAGTG CTTCGGGAGG AAGTATAGTA 2354
GTTACCGGAT TGAAAGTTTG CCACGATTTT CATATTACAT TACATAGATC AATGAAAGAT 2414
AAAATAAGAT TGCATCTTTC TCTTTTATCA AAGGGCATAT TAAAAGATGA AGATCATAAT 2474
AAACTTTCTG GTTATATTGC TTATGCAAAA GATATAGACC CTCATTTTTTA TACAAAACCTG 2534
AACAGAAAAT ATTTTCAAGA AATAAAATGG ATTCAGAATC TCCACAACAA AGTTGAATAA 2594
ACTTTATATT TTGGATGCAC CCAATAACT TCATTGATTA AATTGGGAAC AATATAGGCT 2654
TTTCAGGATG ACCTACACTC TAGAGAATGT GTATACAAAA GTGTATAAGT TATTTTCAAA 2714
CCTATATAAA ATACAGCAAA ATCAATGCAT TGGCGGCATT TTACCACTCC TGTGATCTTC 2774
CGCCAAAATG CCTC 2788

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 316 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met	Arg	Ile	Tyr	Ser	Leu	Ile	Asp	Ser	Gln	Thr	Leu	Met	Thr	Lys	Gly
1				5				10						15	
Phe	Ala	Ser	Glu	Val	Met	Arg	Ser	Pro	Glu	Pro	Pro	Lys	Lys	Trp	Asp
			20					25					30		
Ile	Ala	Lys	Lys	Lys	Gly	Gly	Met	Arg	Thr	Ile	Tyr	His	Pro	Ser	Ser
		35					40					45			

Lys Val Lys Leu Ile Gln Tyr Trp Leu Met Asn Asn Val Phe Ser Lys
 50 55 60
 Leu Pro Met His Asn Ala Ala Tyr Ala Phe Val Lys Asn Arg Ser Ile
 65 70 75 80
 Lys Ser Asn Ala Leu Leu His Ala Glu Ser Lys Asn Lys Tyr Tyr Val
 85 90 95
 Lys Ile Asp Leu Lys Asp Phe Phe Pro Ser Ile Lys Phe Thr Asp Phe
 100 105 110
 Glu Tyr Ala Phe Thr Arg Tyr Arg Asp Arg Ile Glu Phe Thr Thr Glu
 115 120 125
 Tyr Asp Lys Glu Leu Leu Gln Leu Ile Lys Thr Ile Cys Phe Ile Ser
 130 135 140
 Asp Ser Thr Leu Pro Ile Gly Phe Pro Thr Ser Pro Leu Ile Ala Asn
 145 150 155 160
 Phe Val Ala Arg Glu Leu Asp Glu Lys Leu Thr Gln Lys Leu Asn Ala
 165 170 175
 Ile Asp Lys Leu Asn Ala Thr Tyr Thr Arg Tyr Ala Asp Asp Ile Ile
 180 185 190
 Val Ser Thr Asn Met Lys Gly Ala Ser Lys Leu Ile Leu Asp Cys Phe
 195 200 205
 Lys Arg Thr Met Lys Glu Ile Gly Pro Asp Phe Lys Ile Asn Ile Lys
 210 215 220
 Lys Phe Lys Ile Cys Ser Ala Ser Gly Gly Ser Ile Val Val Thr Gly
 225 230 235 240
 Leu Lys Val Cys His Asp Phe His Ile Thr Leu His Arg Ser Met Lys
 245 250 255
 Asp Lys Ile Arg Leu His Leu Ser Leu Leu Ser Lys Gly Ile Leu Lys
 260 265 270
 Asp Glu Asp His Asn Lys Leu Ser Gly Tyr Ile Ala Tyr Ala Lys Asp
 275 280 285
 Ile Asp Pro His Phe Tyr Thr Lys Leu Asn Arg Lys Tyr Phe Gln Glu
 290 295 300
 Ile Lys Trp Ile Gln Asn Leu His Asn Lys Val Glu
 305 310 315

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1602 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 548..1507

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TGGCATCTAT TAAGAAGGTT AGGAAAGAAA ATAAAGTATC AAAAGATATT GGAAATATAT 60
 TATACGCAGA GCGTTTCTAT TGCCTTGTAT CTATTTACTG GATAGTGTCA ACTACCGCAC 120
 ACTGTGTGAA CTAGCTTTTA AAGCGATAAA GCAAGATGAT GTTTTATCTA AAATTATTGT 180
 TAGATCCGTT GTTTCTCGTC TAATAAATGA ACGAAAAATA CTTCAAATGA CTGATGGTTA 240
 TCAGGTCAC TCTTTGGGGG CTAGCTATGT TAGGAGCGTC TTTGATAGAA AGACACTTGA 300
 CCGATTGCGG CTTGAGATTA TGAATTTTGA AAACCGTAGA AAATCAACAT TTAACATGA 360
 TAAGATTCCG TATGCGCACC CTTAGCGAGA GGTTTATCAT TAAGGTCAAC CTCTGGATGT 420
 TGTTTCGGCA TCCTGCATTG AATCTGAGTT ACTGTCTGTT TTCCTTGTTG GAACGGAGAG 480
 CATCGCCTGA TGCTCTCCGA GCCAACCAGG AAACCCGTTT TTTCTGACGT AAGGGTGCGC 540
 AACTTTC ATG AAA TCC GCT GAA TAT TTG AAC ACT TTT AGA TTG AGA AAT 589
 Met Lys Ser Ala Glu Tyr Leu Asn Thr Phe Arg Leu Arg Asn
 1 5 10
 CTC GGC CTA CCT GTC ATG AAC AAT TTG CAT GAC ATG TCT AAG GCG ACT 637
 Leu Gly Leu Pro Val Met Asn Asn Leu His Asp Met Ser Lys Ala Thr
 15 20 25 30
 CGC ATA TCT GTT GAA ACA CTT CGG TTG TTA ATC TAT ACA GCT GAT TTT 685
 Arg Ile Ser Val Glu Thr Leu Arg Leu Leu Ile Tyr Thr Ala Asp Phe
 35 40 45
 CGC TAT AGG ATC TAC ACT GTA GAA AAG AAA GGC CCA GAG AAG AGA ATG 733
 Arg Tyr Arg Ile Tyr Thr Val Glu Lys Lys Gly Pro Glu Lys Arg Met
 50 55 60
 AGA ACC ATT TAC CAA CCT TCT CGA GAA CTT AAA GCC TTA CAA GGA TGG 781
 Arg Thr Ile Tyr Gln Pro Ser Arg Glu Leu Lys Ala Leu Gln Gly Trp
 65 70 75
 GTT CTA CGT AAC ATT TTA GAT AAA CTG TCG TCA TCT CCT TTT TCT ATT 829
 Val Leu Arg Asn Ile Leu Asp Lys Leu Ser Ser Ser Pro Phe Ser Ile
 80 85 90

GGA	TTT	GAA	AAG	CAC	CAA	TCT	ATT	TTG	AAT	AAT	GCT	ACC	CCG	CAT	ATT	877
Gly	Phe	Glu	Lys	His	Gln	Ser	Ile	Leu	Asn	Asn	Ala	Thr	Pro	His	Ile	
95					100					105					110	
GGG	GCA	AAC	TTT	ATA	CTG	AAT	ATT	GAT	TTG	GAG	GAT	TTT	TTC	CCA	AGT	925
Gly	Ala	Asn	Phe	Ile	Leu	Asn	Ile	Asp	Leu	Glu	Asp	Phe	Phe	Pro	Ser	
				115					120					125		
TTA	ACT	GCT	AAC	AAA	GTT	TTT	GGA	GTG	TTC	CAT	TCT	CTT	GGT	TAT	AAT	973
Leu	Thr	Ala	Asn	Lys	Val	Phe	Gly	Val	Phe	His	Ser	Leu	Gly	Tyr	Asn	
			130					135					140			
CGA	CTA	ATA	TCT	TCA	GTT	TTG	ACA	AAA	ATA	TGT	TGT	TAT	AAA	AAT	CTG	1021
Arg	Leu	Ile	Ser	Ser	Val	Leu	Thr	Lys	Ile	Cys	Cys	Tyr	Lys	Asn	Leu	
		145					150					155				
CTA	CCA	CAA	GGT	GCT	CCA	TCA	TCA	CCT	AAA	TTA	GCT	AAT	CTA	ATA	TGT	1069
Leu	Pro	Gln	Gly	Ala	Pro	Ser	Ser	Pro	Lys	Leu	Ala	Asn	Leu	Ile	Cys	
	160					165					170					
TCT	AAA	CTT	GAT	TAT	CGT	ATT	CAG	GGT	TAT	GCA	GGT	AGT	CGG	GGC	TTG	1117
Ser	Lys	Leu	Asp	Tyr	Arg	Ile	Gln	Gly	Tyr	Ala	Gly	Ser	Arg	Gly	Leu	
175					180					185					190	
ATA	TAT	ACG	AGA	TAT	GCC	GAT	GAT	CTC	ACC	TTA	TCT	GCA	CAG	TCT	ATG	1165
Ile	Tyr	Thr	Arg	Tyr	Ala	Asp	Asp	Leu	Thr	Leu	Ser	Ala	Gln	Ser	Met	
				195					200					205		
AAA	AAG	GTT	GTT	AAA	GCA	CGT	GAT	TTT	TTA	TTT	TCT	ATA	ATC	CCA	AGT	1213
Lys	Lys	Val	Val	Lys	Ala	Arg	Asp	Phe	Leu	Phe	Ser	Ile	Ile	Pro	Ser	
			210					215					220			
GAA	GGA	TTG	GTT	ATT	AAC	TCA	AAA	AAA	ACT	TGT	ATT	AGT	GGG	CCT	CGT	1261
Glu	Gly	Leu	Val	Ile	Asn	Ser	Lys	Lys	Thr	Cys	Ile	Ser	Gly	Pro	Arg	
		225					230					235				
AGT	CAG	AGG	AAA	GTT	ACA	GGT	TTA	GTT	ATT	TCA	CAA	GAG	AAA	GTT	GGG	1309
Ser	Gln	Arg	Lys	Val	Thr	Gly	Leu	Val	Ile	Ser	Gln	Glu	Lys	Val	Gly	
	240					245					250					
ATA	GGT	AGA	GAA	AAA	TAT	AAA	GAA	ATT	AGA	GCA	AAG	ATA	CAT	CAT	ATA	1357
Ile	Gly	Arg	Glu	Lys	Tyr	Lys	Glu	Ile	Arg	Ala	Lys	Ile	His	His	Ile	
255					260					265					270	
TTT	TGC	GGT	AAG	TCT	TCT	GAG	ATA	GAA	CAC	GTT	AGG	GGA	TGG	TTG	TCA	1405
Phe	Cys	Gly	Lys	Ser	Ser	Glu	Ile	Glu	His	Val	Arg	Gly	Trp	Leu	Ser	
				275					280					285		
TTT	ATT	TTA	AGT	GTG	GAT	TCA	AAA	AGC	CAT	AGG	AGA	TTA	ATA	ACT	TAT	1453
Phe	Ile	Leu	Ser	Val	Asp	Ser	Lys	Ser	His	Arg	Arg	Leu	Ile	Thr	Tyr	
			290					295					300			
ATT	AGC	AAA	TTA	GAA	AAA	AAA	TAT	GGA	AAG	AAC	CCT	TTA	AAT	AAA	GCG	1501
Ile	Ser	Lys	Leu	Glu	Lys	Lys	Tyr	Gly	Lys	Asn	Pro	Leu	Asn	Lys	Ala	

SECRET FEB 03 1967

315

1557

1602

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1540 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(A) NAME/KEY: CDS

- (B) LOCATION: 396..1352

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ATG GAT GCT ACC CGG ACA
Met Asp Ala Thr Arg Thr
1 5

ACC CTT CTG GCG CTC GAT TTG TTC GGC TCG CCG GGC TGG AGC GCC GAT 461
Thr Leu Leu Ala Leu Asp Leu Phe Gly Ser Pro Gly Trp Ser Ala Asp
10 15 20

AAA GAA ATA CAG CGA CTG CAT GCG CTC AGT AAT CAT GCC GGA CGC CAT 509
Lys Glu Ile Gln Arg Leu His Ala Leu Ser Asn His Ala Gly Arg His
25 30 35

TAC CGA CGC ATT ATT CTT TCT AAA CGC CAC GGT GGT CAG CGG CTG GTG 557
Tyr Arg Arg Ile Ile Leu Ser Lys Arg His Gly Gly Gln Arg Leu Val
40 45 50

TTA GCC CCT GAT TAC TTG CTC AAA ACC GTA CAG CGC AAC ATT CTT AAG 605

Leu Ala Pro Asp Tyr Leu Leu Lys Thr Val Gln Arg Asn Ile Leu Lys
55 60 65 70

AAC GTC CTT TCA CAA TTT CCG CTT TCC CCT TTT GCT ACA GCC TAC CGA 653
Asn Val Leu Ser Gln Phe Pro Leu Ser Pro Phe Ala Thr Ala Tyr Arg
75 80 85

CCA GGT TGC CCA ATC GTC AGC AAC GCG CAG CCA CAC TGC CAA CAG CCG 701
Pro Gly Cys Pro Ile Val Ser Asn Ala Gln Pro His Cys Gln Gln Pro
90 95 100

CAG ATC CTG AAA CTC GAT ATC GAA AAC TTT TTC GAT AGC ATT AGC TGG 749
Gln Ile Leu Lys Leu Asp Ile Glu Asn Phe Phe Asp Ser Ile Ser Trp
105 110 115

TTA CAG GTC TGG CGT GTG TTT CGC CAG GCC CAG TTG CCA CGT AAT GTG 797
Leu Gln Val Trp Arg Val Phe Arg Gln Ala Gln Leu Pro Arg Asn Val
120 125 130

GTA ACC ATG CTG ACC TGG ATT TGT TGT TAT AAC GAC GCG TTA CCG CAG 845
Val Thr Met Leu Thr Trp Ile Cys Cys Tyr Asn Asp Ala Leu Pro Gln
135 140 145 150

GGG GCA CCA ACT TCG CCA GCC ATT TCC AAT CTT GTG ATG CGC CGT TTT 893
Gly Ala Pro Thr Ser Pro Ala Ile Ser Asn Leu Val Met Arg Arg Phe
155 160 165

GAT GAA CGC ATA GGG GAA TGG TGT CAG GCT CGG GGA ATT ACC TAC ACC 941
Asp Glu Arg Ile Gly Glu Trp Cys Gln Ala Arg Gly Ile Thr Tyr Thr
170 175 180

CGC TAC TGC GAT GAC ATG ACC TTT TCA GGT CAC TTC AAT GCC CGC CAG 989
Arg Tyr Cys Asp Asp Met Thr Phe Ser Gly His Phe Asn Ala Arg Gln
185 190 195

GTT AAA AAT AAA GTG TGC GGA TTG TTA GCG GAG CTG GGC CTG AGC CTC 1037
Val Lys Asn Lys Val Cys Gly Leu Leu Ala Glu Leu Gly Leu Ser Leu
200 205 210

AAT AAA CGC AAA GGC TGC CTG ATA GCT GCC TGT AAG CGC CAG CAA GTA 1085
Asn Lys Arg Lys Gly Cys Leu Ile Ala Ala Cys Lys Arg Gln Gln Val
215 220 225 230

ACC GGG ATT GTT GTT AAT CAC AAG CCA CAG CTT GCC CGT GAA GCG CGC 1133
Thr Gly Ile Val Val Asn His Lys Pro Gln Leu Ala Arg Glu Ala Arg
235 240 245

CGG GCG CTG CGT CAG GAG GTG CAT TTG TGC CAA AAA TAT GGC GTT ATT 1181
Arg Ala Leu Arg Gln Glu Val His Leu Cys Gln Lys Tyr Gly Val Ile
250 255 260

TCG CAT CTT AGT CAT CGT GGT GAA CTT GAT CCT TCT GGC GAT CTC CAC 1229
Ser His Leu Ser His Arg Gly Glu Leu Asp Pro Ser Gly Asp Leu His
265 270 275

03603037-03603037

GCA CAG GCA ACG GCG TAT CTT TAT GCT TTG CAG GGA AGA ATA AAC TGG	1277
Ala Gln Ala Thr Ala Tyr Leu Tyr Ala Leu Gln Gly Arg Ile Asn Trp	
280 285 290	
TTA TTG CAA ATC AAC CCT GAG GAT GAG GCC TTT CAA CAG GCG AGA GAG	1325
Leu Leu Gln Ile Asn Pro Glu Asp Glu Ala Phe Gln Gln Ala Arg Glu	
295 300 305 310	
AGT GTA AAG CGA ATG CTG GTT GCA TGG TAAGAAAAGC GTCAGGCAGA	1372
Ser Val Lys Arg Met Leu Val Ala Trp	
315	
CGTTTCTGCC TGACCGTTTA GGGGAGAATT ACTGCAACTG CGCGGCAATT AGCGGCCAGC	1432
GGGCGTCAAA ATCATCCGTC GGGCGGTATT TAAACTCGCT GCGGACAAAA CGTGACAGCA	1492
TACCTTCACA GAAGGCCAGG ATCTGGCTTG CCAGCAGGGT TTCATCGG	1540

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Tyr Xaa Asp Asp
1 4

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Ser Xaa Xaa Xaa
1 4

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Xaa Val Thr Gly
 1 4

255080-1030397